CA/EM (Q12L13)

Total Questions: 20

Most Correct Answers: **#20**

Least Correct Answers: **#7**

1. **Which of these algorithms can be used to fill the missing values**

1/8 (A) KNN for regression

1/8 (B) KNN for classification

4/8 (C) both

0/8 (D) I do not know

2. **Bagging is a technique used to reduce**

2/8 (A) the variance of our predictions

0/8 (B) the bias of our predictions

3/8 (C) both

0/8 (D) I do not know

3. **How can Ensemble methods be constructed?**

0/8 (A) By manipulating the training set

1/8 (B) By manipulating the input features

0/8 (C) By manipulating the class labels

0/8 (D) By manipulating the learning algorithm

2/8 (E) All of them

0/8 (F) None

2/8 (G) I do not know

4. **Repeatedly sampling observations are taken**

2/8 (A) from general population

3/8 (B) original sample data set

1/8 (C) I do not know

0/8 (D) None

5. **Random Forest differs from bagging**

3/8 (A) by a random sample of m predictors

2/8 (B) by bootstrapped training samples

1/8 (C) by adaptive sampling

0/8 (D) I do not know

6. **Boosting differs from bagging**

1/8 (A) by a random sample of m predictors

3/8 (B) by bootstrapped training samples

1/8 (C) by adaptive sampling

0/8 (D) I do not know

7. **Averaging many highly correlated quantities**

2/8 (A) lead to as large of a reduction in variance

0/8 (B) does not lead to as large of a reduction in variance

2/8 (C) lead to as large of a reduction in bias

2/8 (D) I do not know

8. **We can perform a Random forest in R using the function**

3/8 (A) randomForest()

1/8 (B) rf()

0/8 (C) randomF()

0/8 (D) boot()

1/8 (E) I do not know

9. **Random Forest works**

2/8 (A) for classification

0/8 (B) for regression

4/8 (C) both

0/8 (D) I do not know

10. Cluster Analysis is

4/8 A Unsupervised learning technique

3/8 B Supervised learning technique

0/8 C I do not know

11. Distance between records and distance between clusters are the same

0/8 A True

5/8 B False

0/8 C I do not know

12. Which of these is the measure of between clusters distance?

0/8 A Single link

1/8 B Complete link

0/8 C Average link

0/8 D Centroid

4/8 E All of them

1/8 F I do not know

13. Single link is

2/8 A the smallest distance between an element in one cluster and an element in the other

2/8 B the largest distance between an element in one cluster and an element in the other

0/8 C the average distance between an element in one cluster and an element in the other

0/8 D distance between the centroids of two clusters

2/8 E I do not know

14. Complete link is

3/8 A the smallest distance between an element in one cluster and an element in the other

2/8 B the largest distance between an element in one cluster and an element in the other

0/8 C the average distance between an element in one cluster and an element in the other

0/8 D distance between the centroids of two clusters

2/8 E I do not know

15. Which of these is the nested algorithm of clustering?

5/8 (A) Hierarchical clustering

0/8 (B) k-means

1/8 (C) Knn

0/8 (D) I do not know

16. Which of these is the unnested algorithm of clustering?

1/8 (A) Hierarchical clustering

4/8 (B) k-means

0/8 (C) Knn

0/8 (D) I do not know

17. Which of these is the type of hierarchical clustering?

1/8 (A) Agglomerative Methods

0/8 (B) Divisive Methods

3/8 (C) Both

1/8 (D) I do not know

18. This function can be used to perform hierarchical clustering in R

4/8 (A) hclust()

1/8 (B) cluster()

0/8 (C) hierarchical ()

0/8 (D) I do not know

19. This function can be used to perform k-means clustering in R

3/8 (A) kmeans()

2/8 (B) kclust()

0/8 (C) kmenscl()

1/8 (D) I do not know

20. Do we need to worry about scaling in clustering?

5/8 A Yes

0/8 B No

0/8 C I do not know